PSC 202
SYRACUSE UNIVERSITY

INTRODUCTION TO POLITICAL ANALYSIS

HYPOTHESIS TESTING WHEN USING SAMPLES, HYPOTHESIS TESTING WITH ONE CONFOUNDER

HOUSEKEEPING

- No in-person sections on Friday
- Instead, we will distribute a worksheet to complete at your leisure
  - Due December 1 (Friday in 2 weeks)
  - Graded pass/fail, counts towards section attendance/participation
- If you have questions about the material, please email and/or attend student hours

TODAY

- Finishing up hypothesis testing with a sample
- Hypothesis testing with one confounder

IDEA

- We start out thinking H_0 is true
  - No relationship between independent and dependent variable in population
- We have a sample that shows a relation difference
  - Do we reject H_0?
  - If we do, we want to do so wrongly at most 5% of time

IDEA

- Ask: If H_0 is true (no difference in population), what is the probability (p) of observing a relation as large (or greater) as we did in our sample *just by chance*?
- If less than 5% ($p<0.05$): we reject H_0
- If more than 5% ($p>0.05$): we don't reject H_0

NOW

- How exactly do we do this hypothesis testing?
  - How do we compute a p-value, etc.?

IN OUR CASE

Job Approval Ratings of President Biden, by Subgroup

	Approve	Disapprove	N
	%	%	
All U.S. adults	56	39	2,937
Gender			
Men	49	45	1,643
Women	62	34	1,294

- H_0: No difference between men and women in population
- The survey does find a difference of 13 percentage points
  - 62 for women vs. 49 for men
  - Instead of 13 percentage points, we use 0.13

IN OUR CASE

Job Approval Ratings of President Biden, by Subgroup

	Approve %	Disapprove %	N
All U.S. adults	56	39	2,937
Gender			
Men	49	45	1,643
Women	62	34	1,294

- Question: If there is no difference between men and women in the population, what is the probability of getting a *sample* where they are at least 13 points different from each other *just by chance*?
  - Specifically: is it lower than 5%?

IN OUR CASE

Job Approval Ratings of President Biden, by Subgroup

	Approve	Disapprove	N
	%	%	
All U.S. adults	56	39	2,937
Gender			
Men	49	45	1,643
Women	62	34	1,294

- Equivalent: If we reject H_0 based on this survey, what is probability of committing Type I error?
  - And is it lower than 5%?

TEST STATISTIC

- **Test statistic t:**

$$t = \frac{H_A - H_0}{\text{Standard Error of Difference}}$$

- **H_A: observed difference between samples**
  - **here: 0.13 (13 percentage points)**
- **H_0: difference between samples if H_0 is true (0.00)**
- **Standard Error of Difference between the two samples (here 0.018)**
  - **I calculated this for you**

TEST STATISTIC

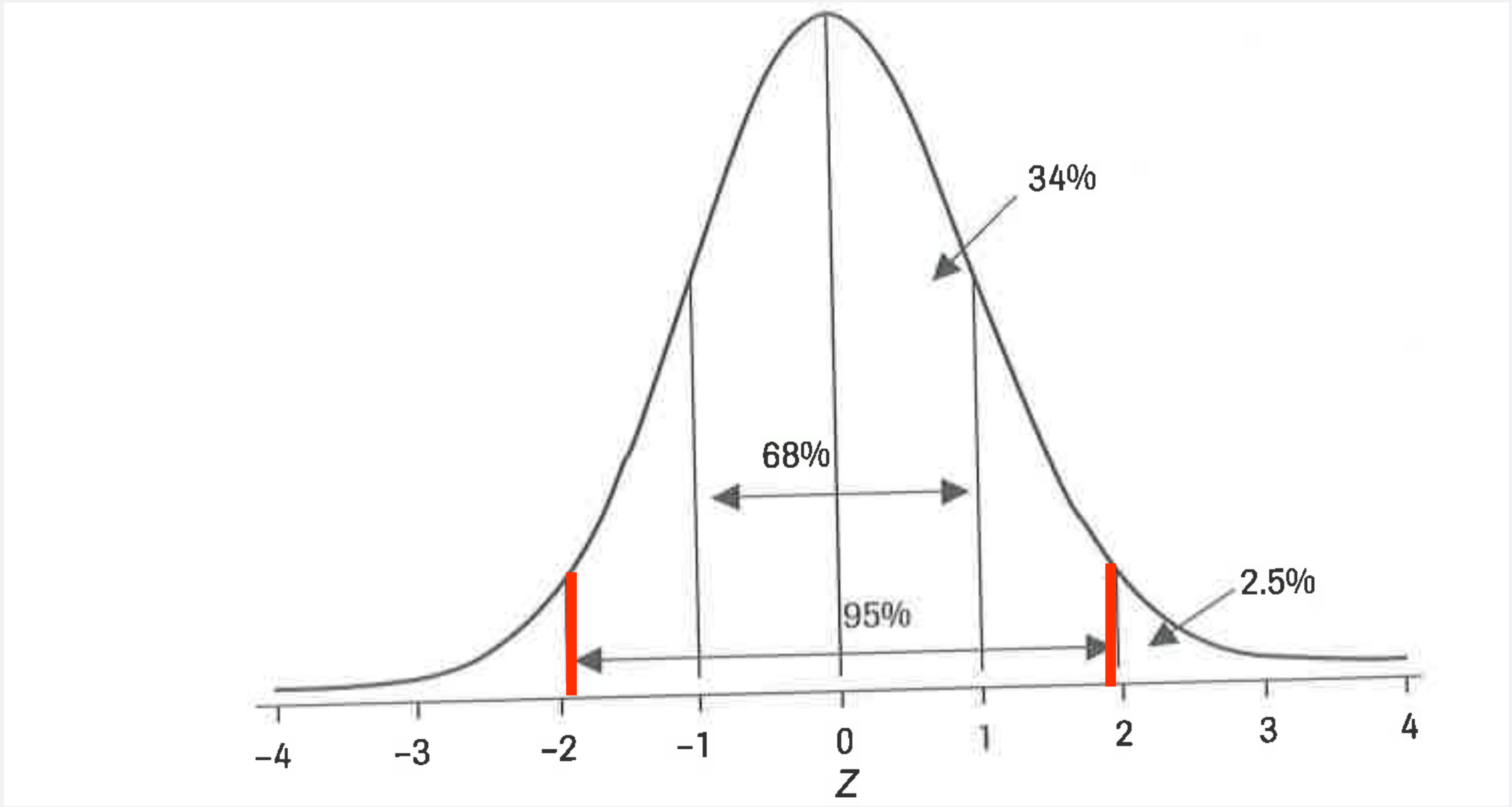
- H_A: 0.13
- H_0: 0
- Standard Error of Difference: 0.018

$$t = \frac{H_A - H_0}{\text{Standard Error of Difference}}$$

$$t = \frac{0.13 - 0.00}{0.018} = 7.22$$

- **This is called the "t-statistic" or "t-ratio"**

NORMAL DISTRIBUTION

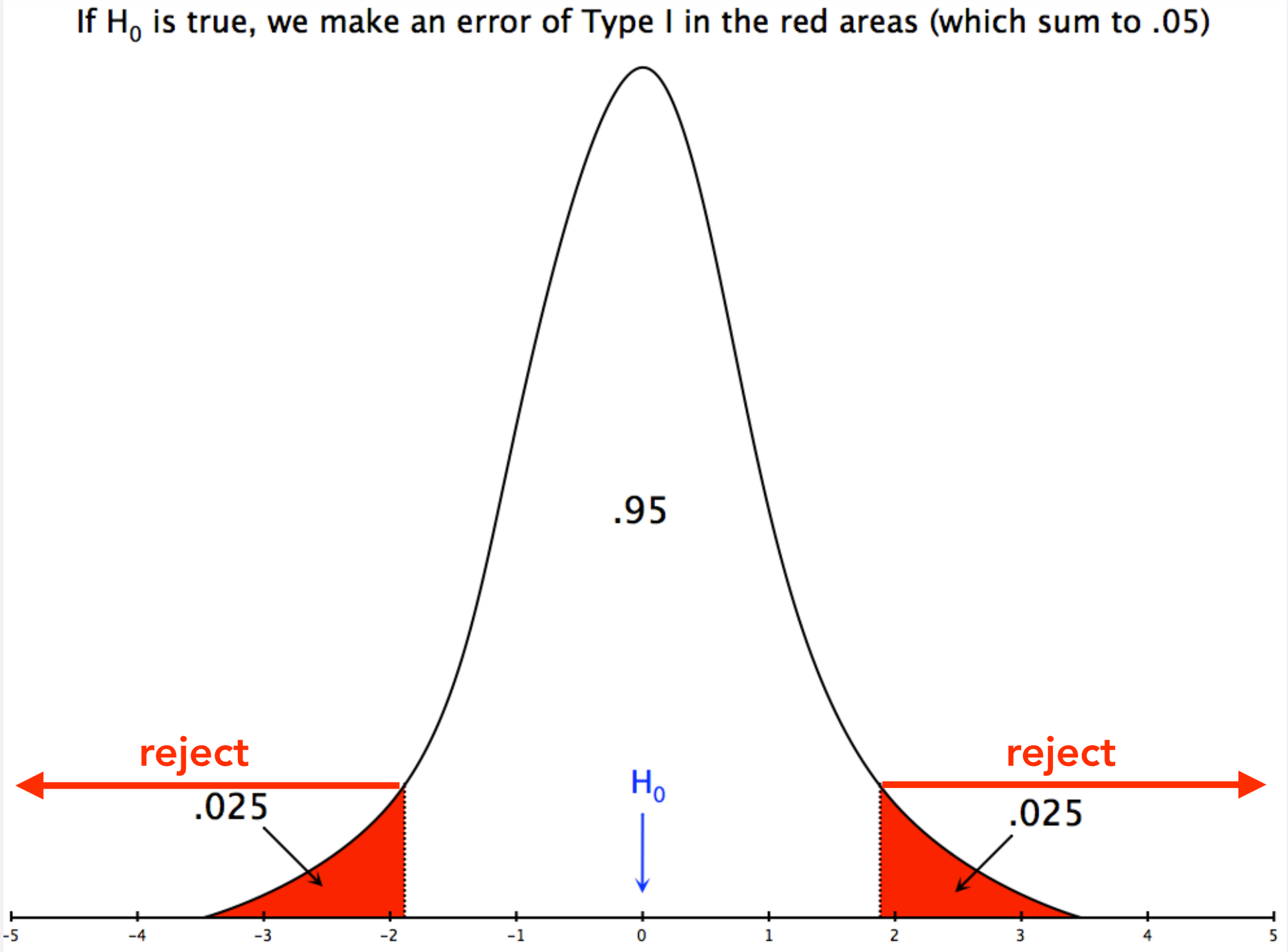


- Remember: 95% between scores of -1.96 and 1.96
- 5% of scores outside of those scores
- T-statistic is (basically) normally distributed

SIGNIFICANCE TEST

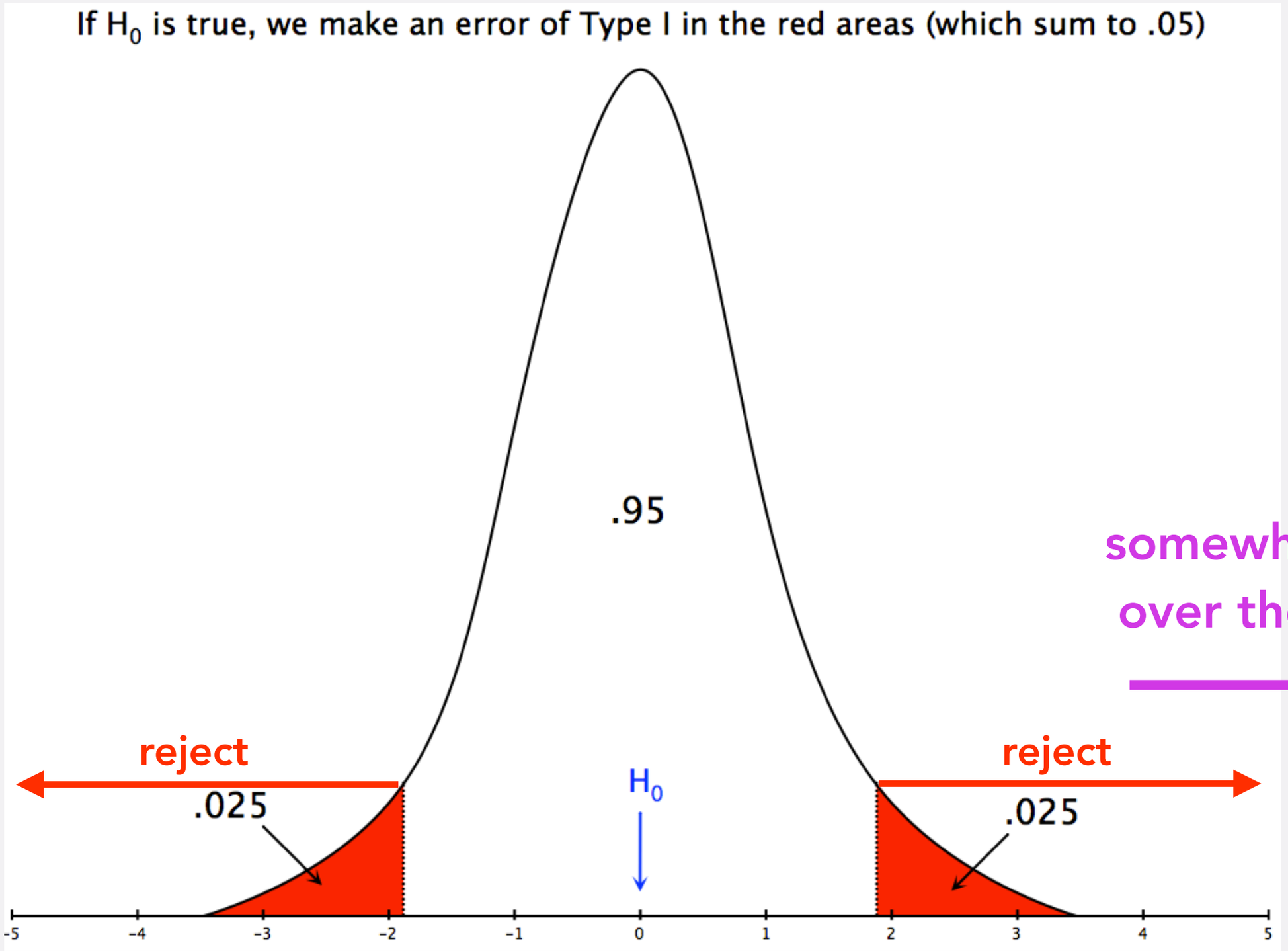
- **We reject H_0 (no difference between men and women) if t-value indicates that chance that we commit a Type I error is less than 5%**
- **5% chance that we falsely reject H_0**

SIGNIFICANCE TEST



- We reject H_0 if t<-1.96 or t>1.96

SIGNIFICANCE TEST



- **t-score: 7.22**

SIGNIFICANCE TEST

Job Approval Ratings of President Biden, by Subgroup

	Approve %	Disapprove %	N
All U.S. adults	56	39	2,937
Gender			
Men	49	45	1,643
Women	62	34	1,294

- If there is no difference between men and women in population, chance that we find 13 percentage points difference in a random sample just by chance is less than 5 percent

SIGNIFICANCE TEST

Job Approval Ratings of President Biden, by Subgroup

	Approve %	Disapprove %	N
All U.S. adults	56	39	2,937
Gender			
Men	49	45	1,643
Women	62	34	1,294

- So we reject the null hypothesis that there is no difference between men and women in approval of Biden
- In favor of the alternative hypothesis that there is a gender difference

ANOTHER EXAMPLE

- From the class survey:
- How would you say the economy is doing?
  - Bad or very bad: 48%
  - Neither, good, very good: 52%

PARTISANSHIP AND ECONOMY

	Democrat	Not Democrat	Total
Bad Or Very Bad	45% (25)	53% (17)	48% (42)
Neither, Good, Or Very Good	55% (30)	47% (15)	52% (45)
Total	100% (55)	100% (32)	100% (87)

- Difference: 8% (0.08)

CROSS-TABULATION

- **Difference between Democrats and non-Democrats is 0.08 (8%)**
  - **Standard error of difference: 0.11**

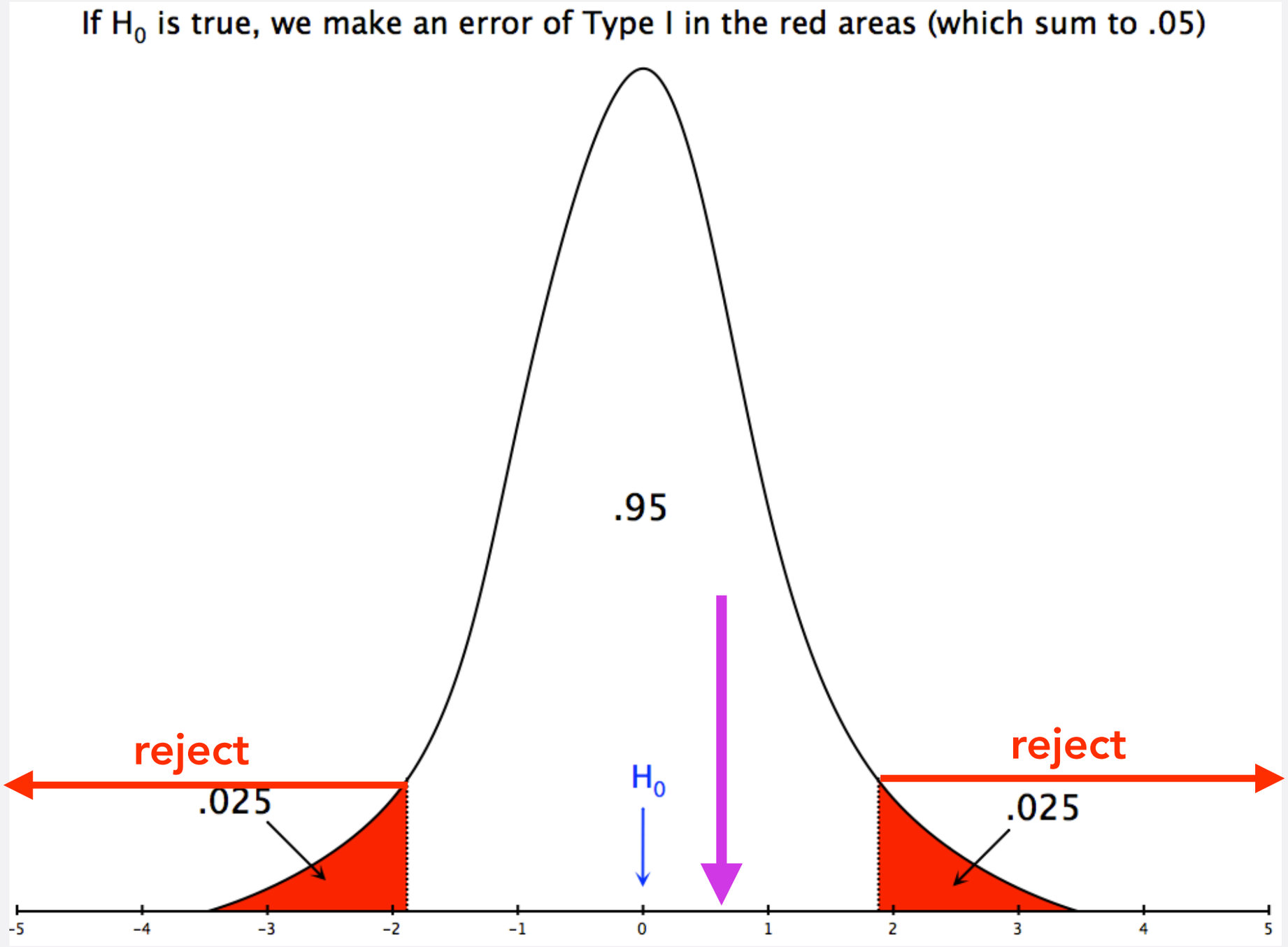
$$\frac{H_A - H_0}{\text{Standard Error of Difference}}$$

$$= \frac{0.08 - 0.0}{0.11}$$

$$= 0.73$$

- **Is this t-statistic large enough to reject H_0?**

SIGNIFICANCE TEST



- We reject H_0 if t<-1.96 or t>1.96
- We had: t=0.73

REJECT H_0?

- We reject H_0 if t < -1.96 or t > 1.96
- We had t = 0.73
- So we cannot reject H_0 that there is no difference between Democrats and non-Democrats in perceptions of economy

SIGNIFICANCE TEST

- If there is no difference in perceptions of economy between Democrats and non-Democrats in population, it is quite likely that we see a difference of 8 percentage points (or larger) in a random sample *just by chance*
- The probability of this happening is larger than 5%

BIVARIATE RELATIONSHIPS

Dependent Variable	Independent Variable: Nominal/Ordinal	Independent Variable: Interval
Nominal/Ordinal	Cross-Tabulation	Not In This Class…
Interval	Mean Comparison	Correlation Coefficient, Linear Regression

BIVARIATE RELATIONSHIPS

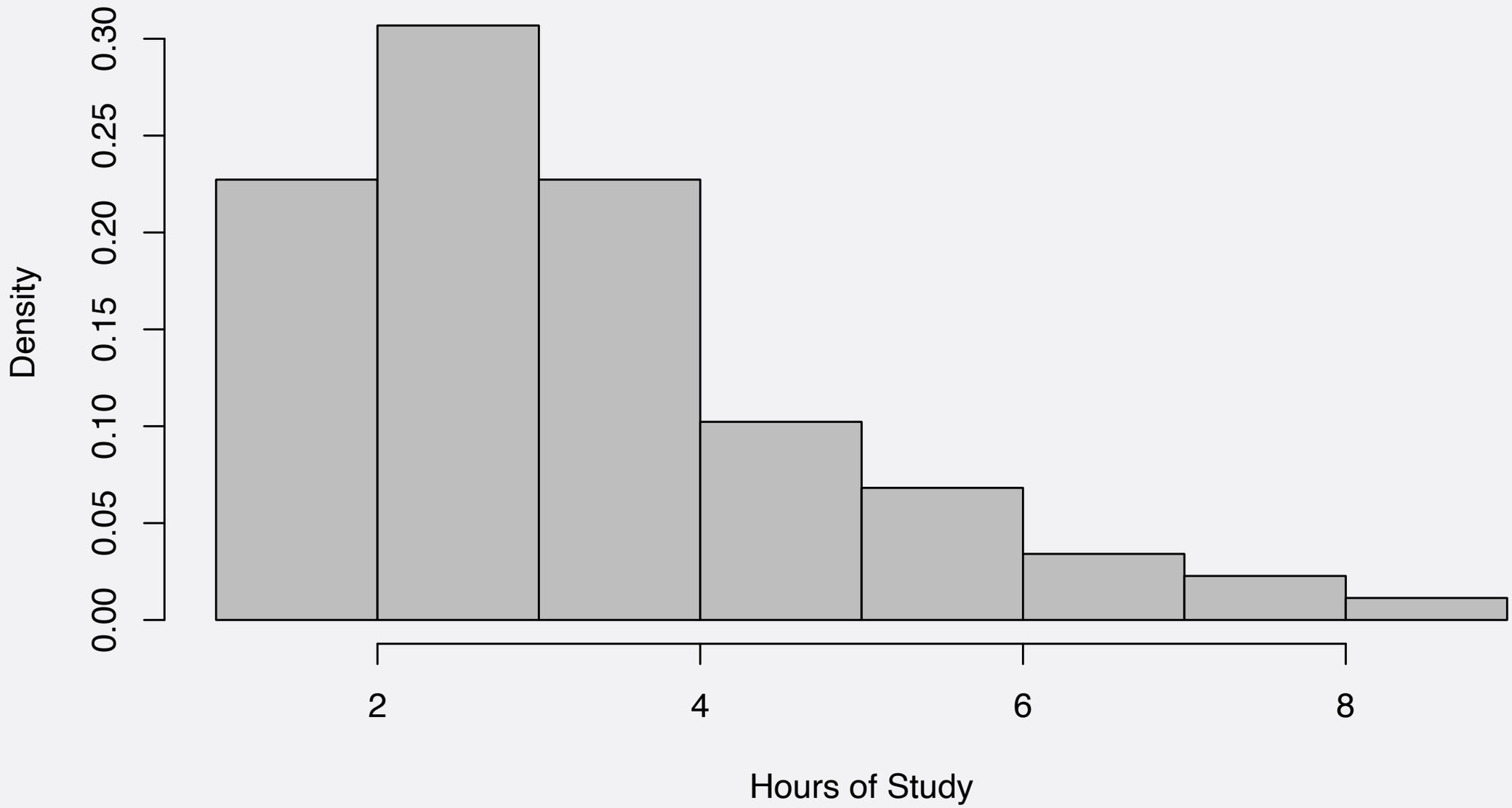
Dependent Variable	Independent Variable: Nominal/Ordinal	Independent Variable: Interval
Nominal/Ordinal	Cross-Tabulation	Not In This Class…
Interval	Mean Comparison	Correlation Coefficient, Linear Regression

CROSS-TABULATION

- Very similar approach as for mean comparisons

EXAMPLE

- On a typical day, how many hours do you spend studying/ revising/preparing for your classes, not counting time in class itself?



GENDER AND STUDYING

Gender	Mean Hours	Frequency	Standard Error
Female	3.68	59	0.21
Male	3.14	31	0.27

GENDER AND STUDYING

Gender	Mean Hours	Frequency	Standard Error
Female	3.68	59	0.21
Male	3.14	31	0.27
Difference	0.54	90	0.34

- Do men really study less than women?

TEST STATISTIC

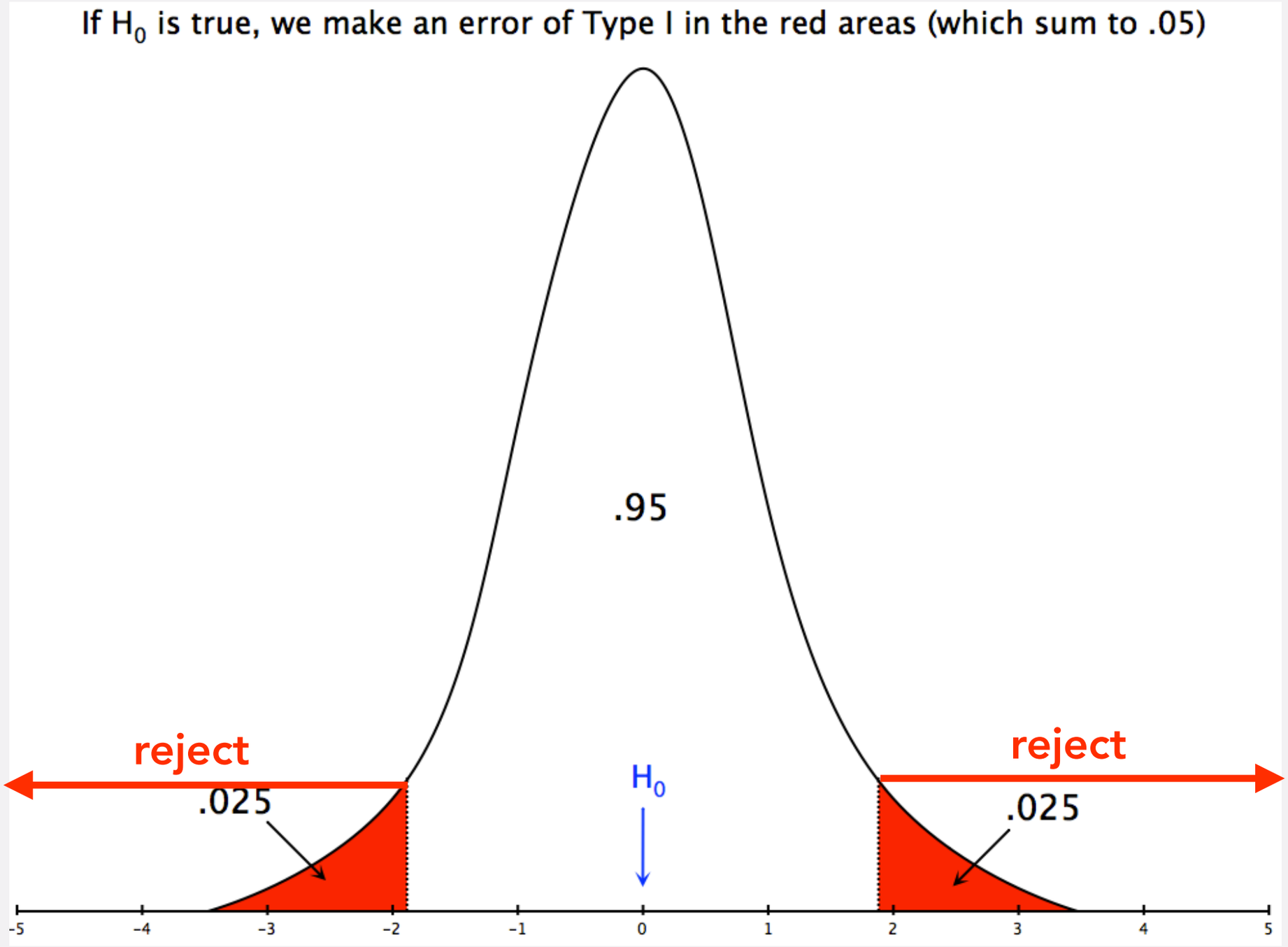
- H_A: 0.54
- H_0: 0
- Standard Error of Difference: 0.34

$$t = \frac{H_A - H_0}{\text{Standard Error}}$$

$$= \frac{0.54 - 0.0}{0.34}$$

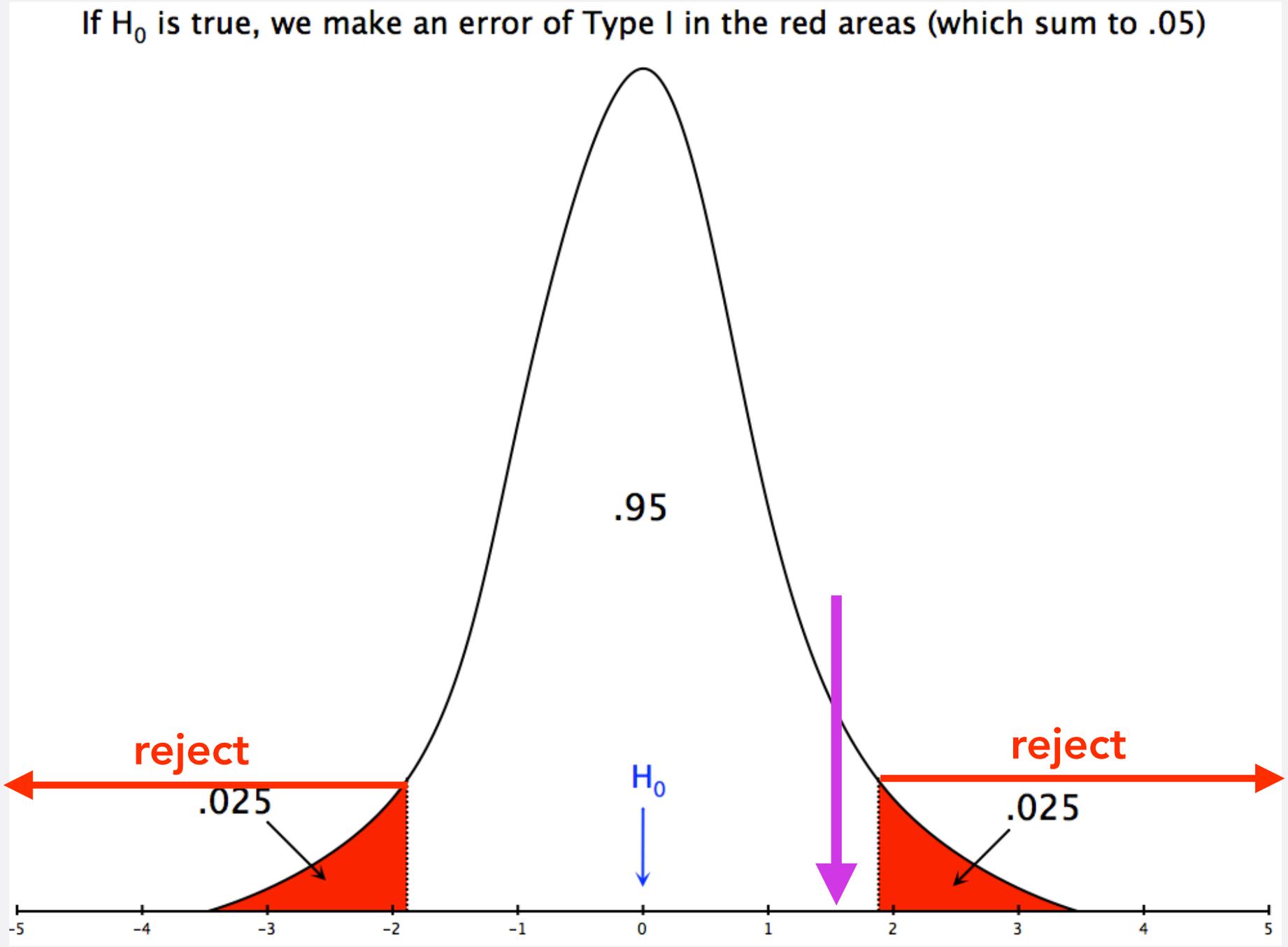
$$= 1.59$$

SIGNIFICANCE TEST



- We reject H_0 if t<-1.96 or t>1.96
- This is equivalent to p<0.05

SIGNIFICANCE TEST



- t-score: 1.59

SIGNIFICANCE TEST

- We cannot reject H_0
- If there is no difference in study time between men and women in population of students, it is quite likely that we see a difference of 0.54 hours (or larger) in a sample of 90 students *just by chance*
- The probability of this happening is larger than 5%

EXERCISE

- Survey: ANES 2016
- DV: Opinion about Obamacare
- 1=favor a great deal, 7=oppose a great deal
- mean=4.09
- n=1,606

EXERCISE

Partisanhsip	Mean Evaluation	Frequency
Dem	2.92	924
Rep	5.69	682
Difference	2.77	1606

- Standard Error of Difference: 0.098

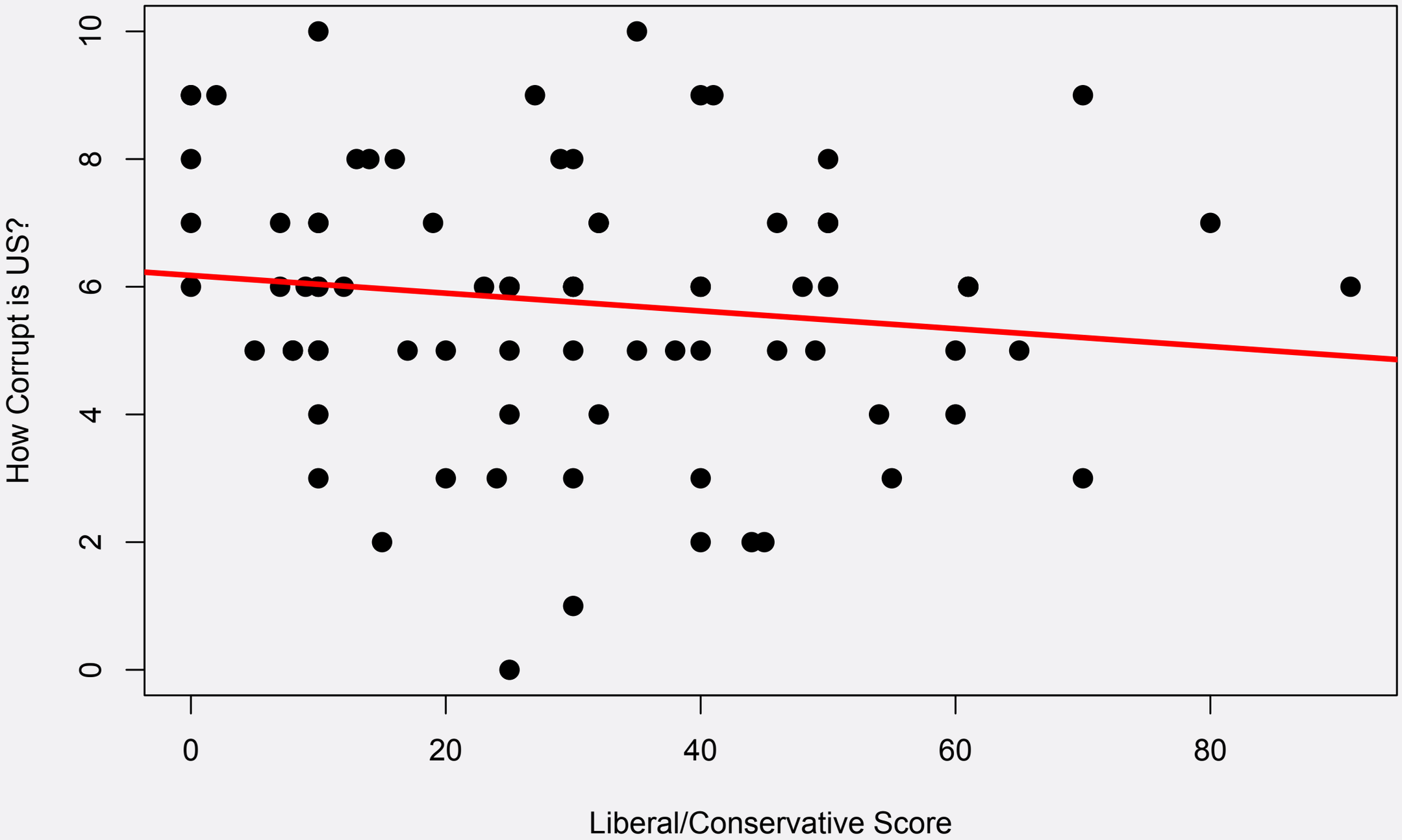
EXERCISE

- Calculate t-statistic and decide whether we can reject H_0
- Solution on last slide (don’t peek)

BIVARIATE RELATIONSHIPS

Dependent Variable	Independent Variable: Nominal/Ordinal	Independent Variable: Interval
Nominal/Ordinal	Cross-Tabulation	Not In This Class…
Interval	Mean Comparison	Correlation Coefficient, Linear Regression

REGRESSION



- **Corruption Score = 6.2 - 0.014 * Lib/Cons**

REJECT H_0?

- Can we reject H_0 that there is no relationship between lib/cons and perceptions of corruption?

FORMULA

$$t = \frac{H_A - H_0}{\text{Standard Error}}$$

- H_A: -0.014
- H_0: 0
- Here, the relevant standard error is the SE of the linear regression coefficient

REGRESSION TABLE

```
> m <- lm(corruption_1 ~ libcons_1, data = data)
> summary(m)

Call:
lm(formula = corruption_1 ~ libcons_1, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-5.8297 -1.0663  0.1424  1.2677  4.3095

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.17768    0.41500  14.886   <2e-16 ***
libcons_1   -0.01392    0.01126  -1.236     0.22
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

REJECT H_0?

$$t = \frac{H_A - H_0}{\text{Standard Error}}$$

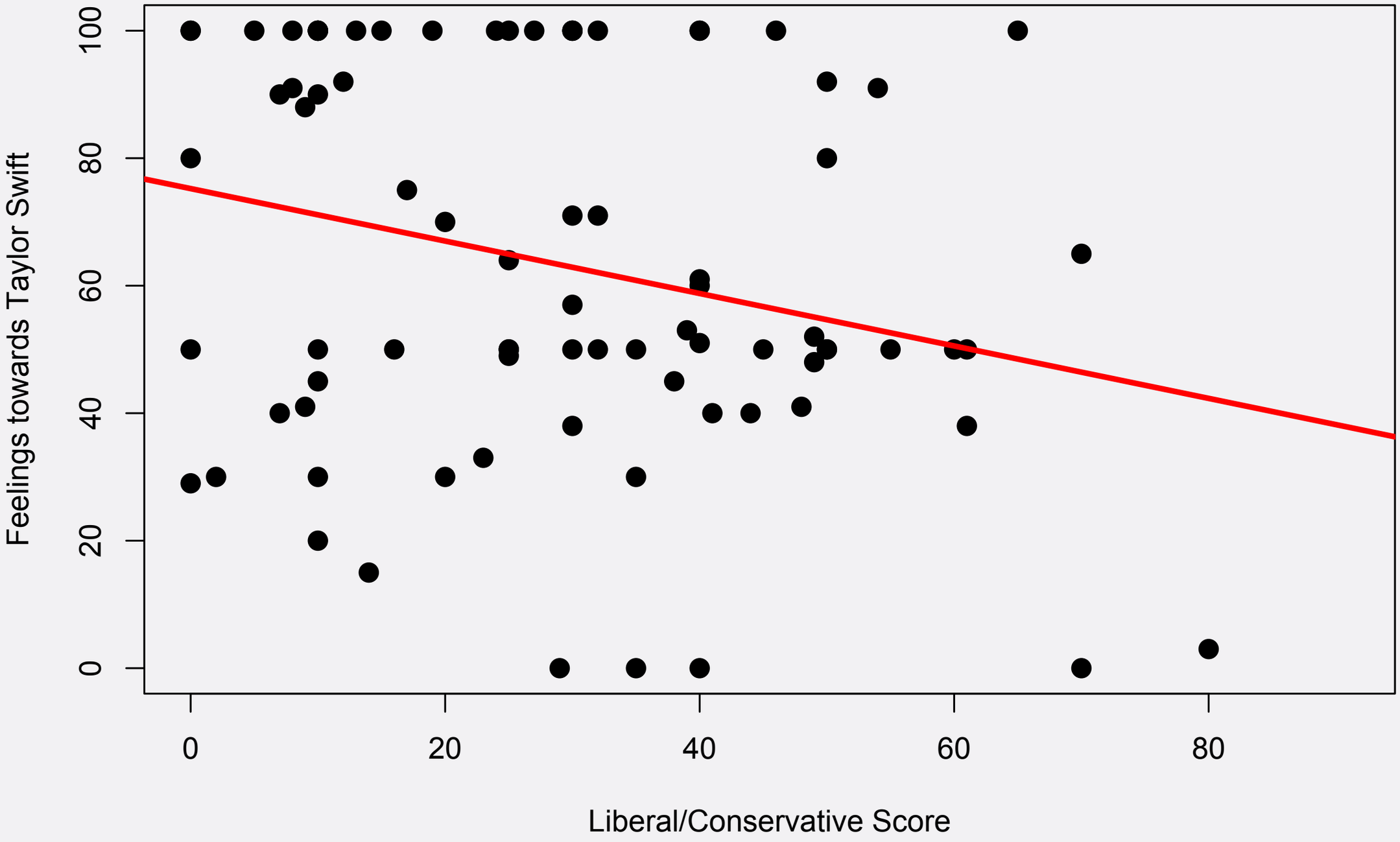
$$= \frac{-0.014 - 0}{0.011}$$

$$= -1.273$$

SIGNIFICANCE TEST

- If how liberal/conservative people are has no effect on corruption perceptions in population, it is quite likely that in a random sample we would see a slope coefficient of -0.014 (or larger) just by chance
  - The probability of this happening is larger than 5%
- We do not reject H_0 and maintain that there is no relation between ideology and corruption perceptions

ANOTHER EXAMPLE



- **Feeling Thermometer = 75.2 - 0.41 * Lib/Cons**

REGRESSION TABLE

```
> m <- lm(therm_6 ~ libcons_1, data = data)
> summary(m)

Call:
lm(formula = therm_6 ~ libcons_1, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-63.301 -18.364  -0.946  28.061  51.509

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  75.2304     5.9645  12.613   <2e-16 ***
libcons_1    -0.4114     0.1720  -2.392   0.0191 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

REJECT H_0?

$$t = \frac{H_A - H_0}{\text{Standard Error}}$$

$$= \frac{-0.41 - 0}{0.17}$$

$$= 2.41$$

SIGNIFICANCE TEST

- If how liberal/conservative people are has no effect on feelings about T. Swift in population, it is quite unlikely that in a random sample, we would see a slope coefficient of -0.41 just by chance
  - The probability of this happening is smaller than 5%
- So we are feel comfortable to reject H_0 and instead conclude that there is a relation between ideology and feelings towards Swift

RECAP

- We are now able to…
  - …tell whether there is covariation between X and Y in a sample
  - …tell whether our evidence (from a sample) is strong enough to conclude with reasonable certainty that the covariation is also present in the population

NEXT STEP

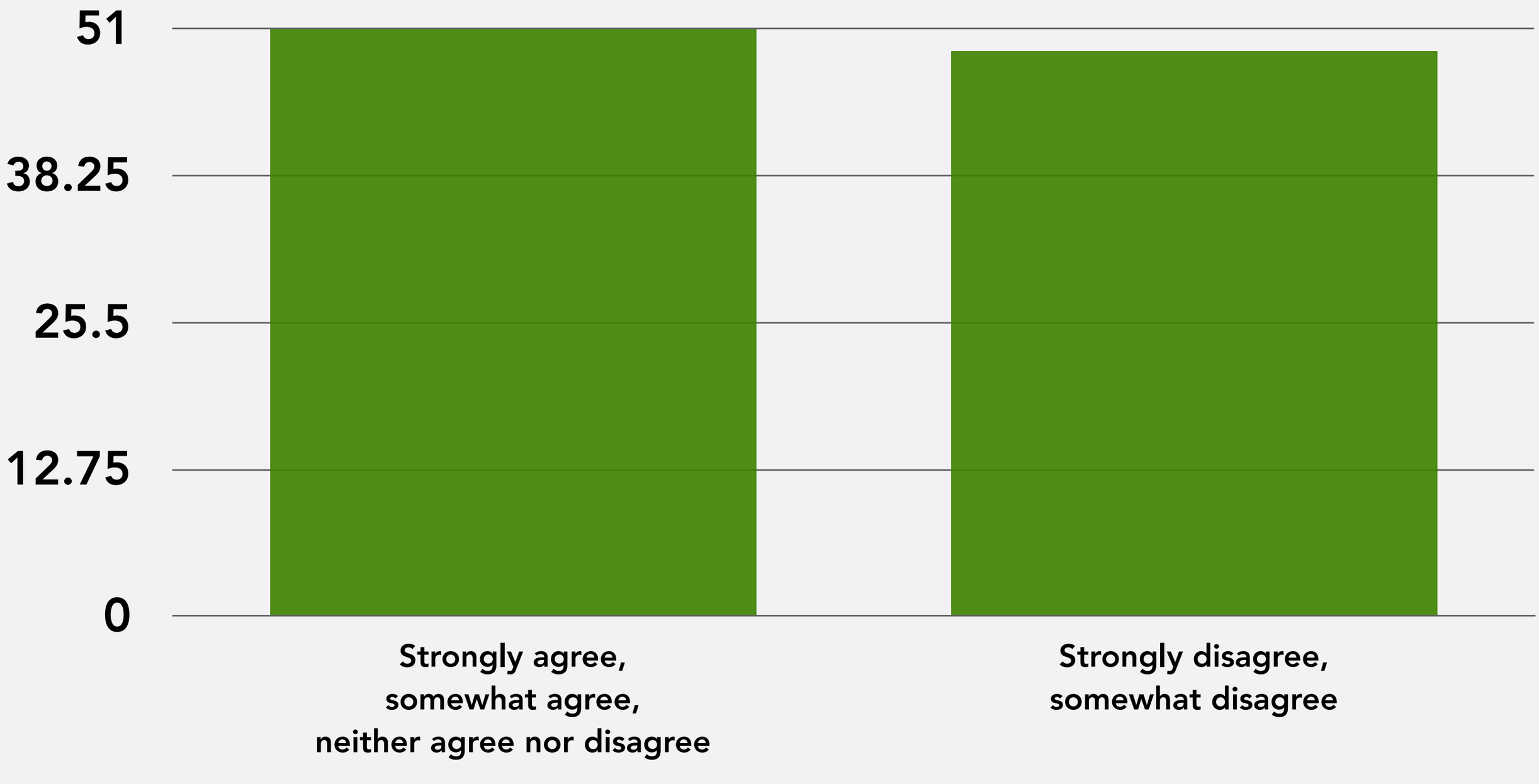
- Is there a credible causal mechanism that connects X to Y?
- Can we rule out the possibility that Y could cause X?
- Is there covariation between X and Y?
- Have we controlled for all confounding variables (Z) that might make the association between X and Y spurious?

TODAY

- Finishing up hypothesis testing with a sample
- Hypothesis testing with one confounder

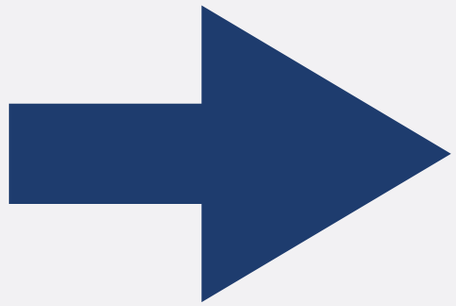
SURVEY

- How much do you agree with the following statement: I would feel safer if there was more armed security personnel on campus.



BIVARIATE RELATIONSHIP

?



Feeling safer if more armed security

- What explains why some of you would feel safer with more armed security on campus, while others would not feel safer?

PARTISANSHIP & SAFETY

	Democrats	Not Democrats	Total
Feel Safer	48% (27)	56% (18)	51% (45)
Not Feel Safer	52% (29)	44% (14)	49% (43)
Total	100% (56)	100% (32)	100% (88)

BIVARIATE RELATIONSHIP

Partisanship → Feeling safer if more armed security

- Zero-order effect: Non-Democrats are 8 percentage points more likely to feel safer with more armed security than Democrats

CAUSALITY

- Want to know *causal* effect of partisanship on feeling safer with armed security:
- Feeling of person if Democrat - Feeling of same person if not Democrat
- For each person, only one of those is observed
- Fundamental problem of causal inference: We can't observe alternate reality in which you identify with a different party!

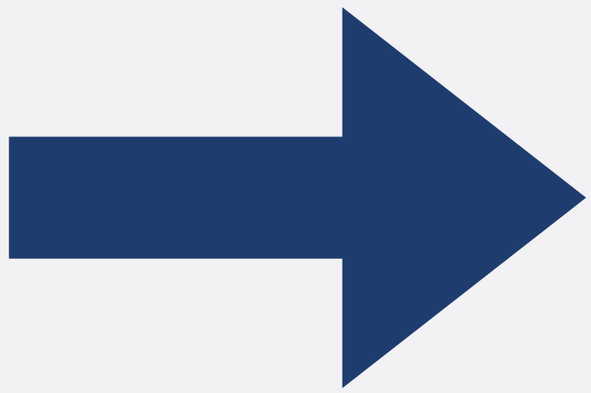
CAUSALITY

- What we can compute:
- Feeling of people who are Democrats - Feeling of people who are not Democrat
- Problem: Students who *choose* to identify as Democrats are likely different from students who *choose* to not identify as Democrats in many other ways
- These other differences potentially affect our ability to compute the causal effect of partisanship

CONFOUNDER?

Race (Z)

Partisanship (X)



Feeling safer if more armed security (Y)

MAYBE THIS IS GOING ON?

Non-white students more likely to be Democrats than white students

Race (Z)

Partisanship (X)

Feeling safer if more armed security (Y)

MAYBE THIS IS GOING ON?

Non-white students more likely to be Democrats than white students

Race (Z)

Non-white students less likely to feel safer with armed security than white students

Partisanship (X)

Feeling safer if more armed security (Y)

MAYBE THIS IS GOING ON?

Non-white students more likely to be Democrats than white students

Race (Z)

Non-white students more likely to not feel safer with armed security than white students

Partisanship (X)

Feeling safer if more armed security (Y)

Partisanship by itself has no effect on feeling safer

POTENTIAL CONCERN

Disproportionately non-white students

Disproportionately white students

	Democrats	Not Democrats	Total
Feel Safer	48% (27)	56% (18)	51% (45)
Not Feel Safer	52% (29)	44% (14)	49% (43)
Total	100% (56)	100% (32)	100% (88)

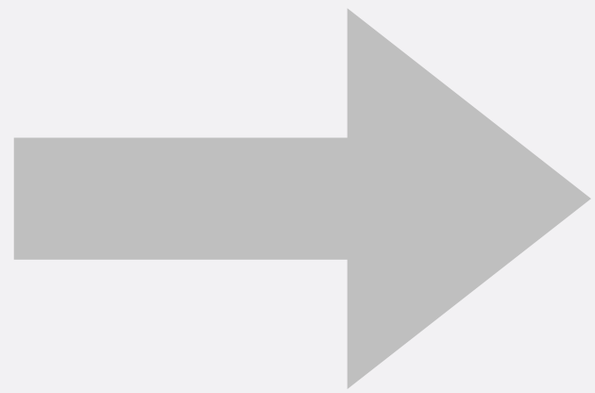
MAYBE THIS IS GOING ON?

Non-white students more likely to be Democrats than white students

Race (Z)

Non-white students more likely to not feel safer with armed security than white students

Partisanship (X)



Feeling safer if more armed security (Y)

- How can we find out if this is what's going on?

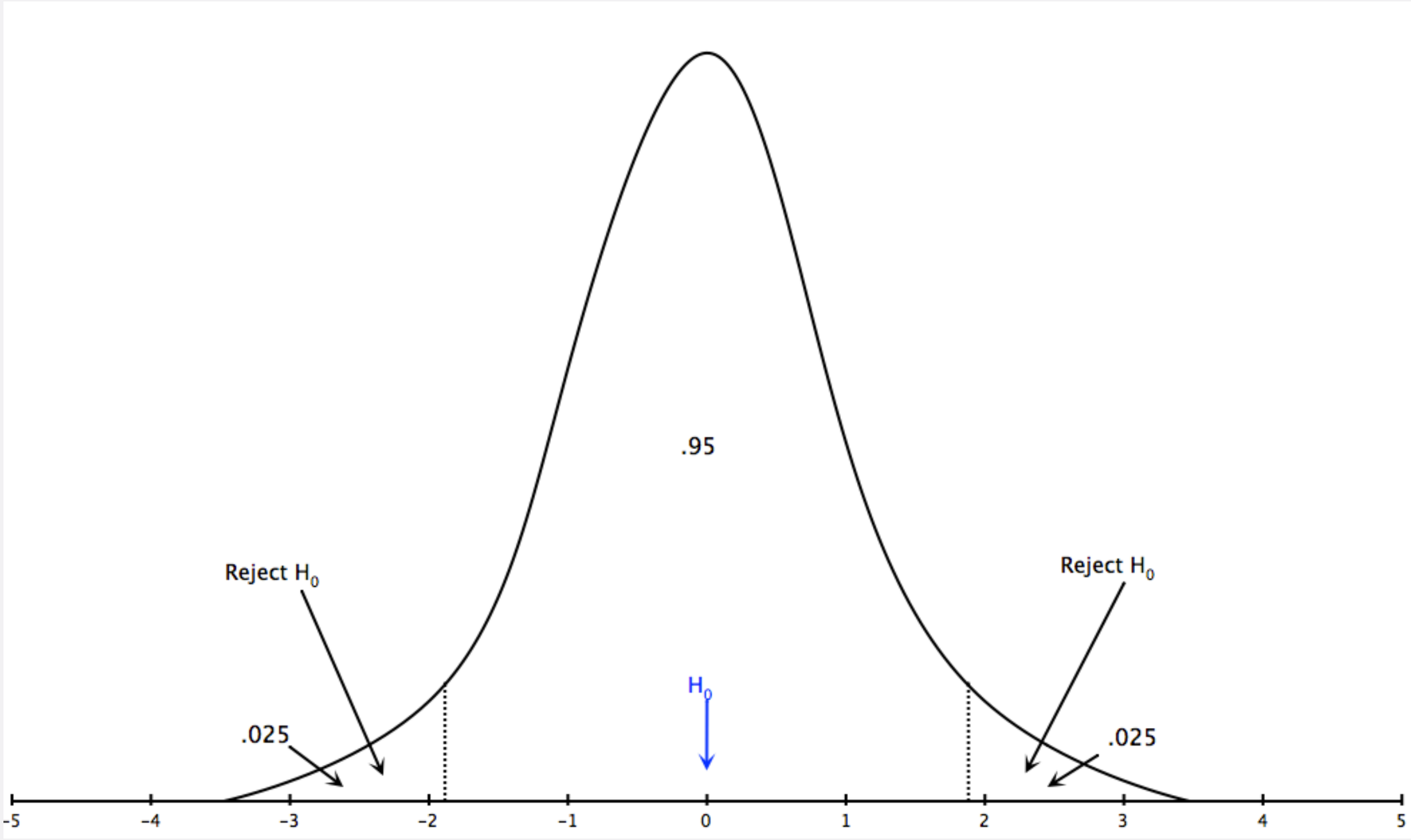
EXERCISE SOLUTION

- H_A: 2.77
- H_0: 0
- Standard Error of Difference: 0.098

$$t = \frac{H_A - H_0}{\text{Standard Error of Difference}}$$

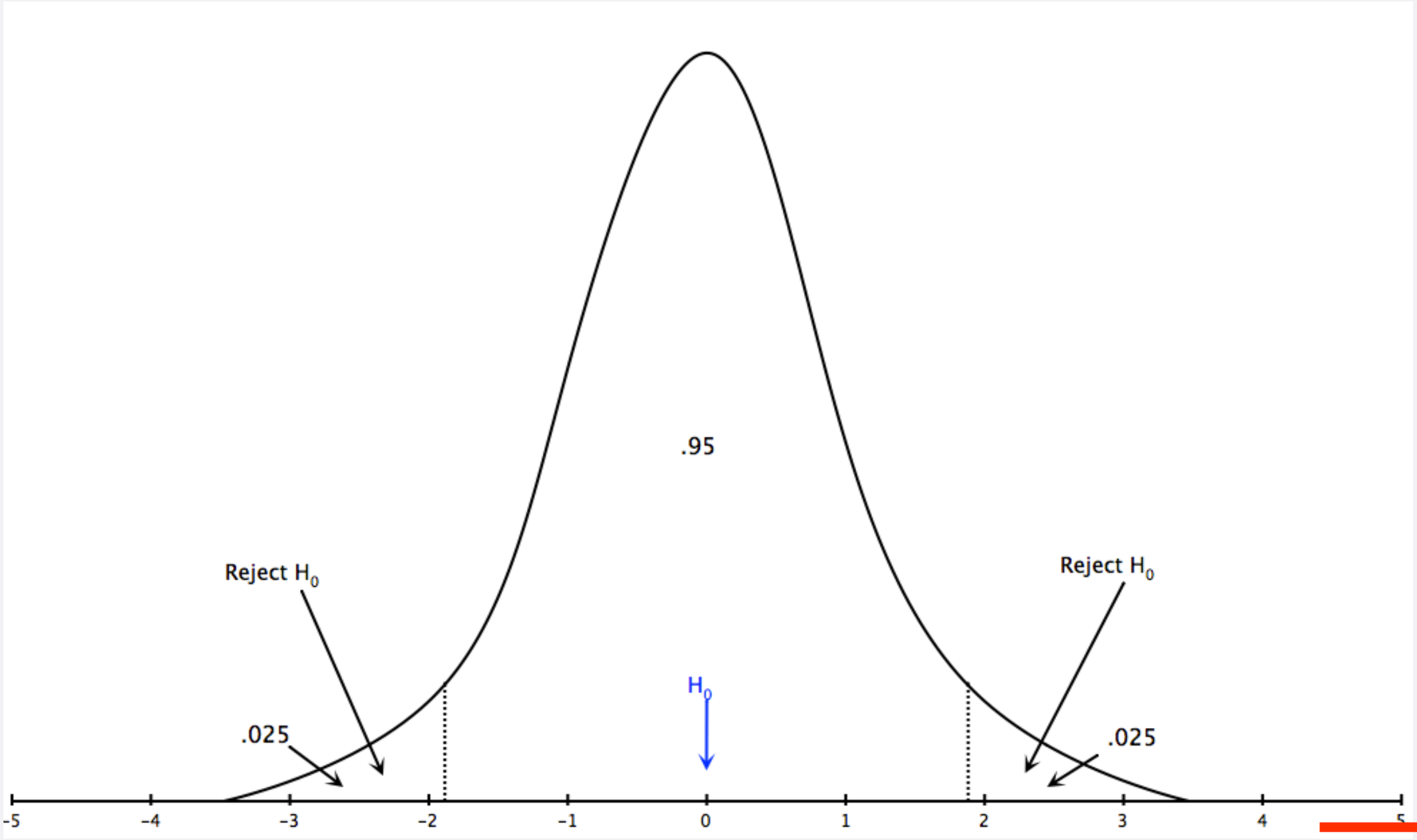
$$t = \frac{2.77 - 0.00}{0.098} = 28.26$$

EXERCISE SOLUTION



- We reject H_0 if t<-1.96 or t>1.96
- This is equivalent to p<0.05

EXERCISE SOLUTION



way over there

- t-score: 28.26

EXERCISE SOLUTION

- With n=1,606, a mean difference of 2.77 (SE 0.098) produces a t-statistic of 28.26
- We reject H_0 if t < -1.96 or t > 1.96
  - If there is no difference in the population, it is extremely unlikely to find a large difference of 2.77 points (or larger) in such a large sample just *by chance*
- We reject null hypothesis that there is no difference between R and D in evaluation of Obamacare